Hamburger Beiträge
zur Angewandten Mathematik

Parameter Estimation for Exponential Sums from Sparse Frequency Projections

Benedikt Diederichs, Armin Iske

Nr. 2016-05
March 2016

Preprint submitted to
SAMPLING THEORY IN SIGNAL AND IMAGE PROCESSING

DVI file produced on
March 14, 2016

Parameter Estimation for Exponential Sums from Sparse Frequency Projections

Benedikt Diederichs
Department of Mathematics, University of Hamburg
Bundesstraße 55, 20146 Hamburg, Germany
benedikt.diederichs@uni-hamburg.de

Armin Iske
Department of Mathematics, University of Hamburg
Bundesstraße 55, 20146 Hamburg, Germany
armin.iske@uni-hamburg.de

Abstract

The determination of frequencies that occur in a signal is regarded as a challenging problem in information processing. One particular instance is that of parameter estimation for bivariate exponential sums. Recently proposed methods solve this problem by the application of one-dimensional reconstruction schemes from frequency projections along only a few lines. In this case, sparse samples of projections from the unknown frequency vectors are taken. This, however, may lead to undesired cancellations in the reconstruction of the frequencies. In this paper, we show how to tackle this problem and we discuss how many lines are necessary for reliable reconstructions in such critical situations.

Key words and phrases : Parameter estimation, bivariate exponential sums, sparse recovery, frequency reconstruction, Prony's method

2000 AMS Mathematics Subject Classification — 65T40, 94A12, 94A20

1 Introduction

Exponential sums model signals which are sparse in the frequency domain. Often, one wishes to recover the unknown frequencies of such signals from only a few samples. In the special case of one dimension, i.e., for univariate exponential sums, the reconstruction is well understood and a vast literature is available. Moreover, a variety of numerical algorithms have been developed, including APM [9], ESPRIT [13], and matrix pencil methods [14].

In the special case of two dimensions, one wishes to estimate the parameters of a bivariate exponential sum

$$f(\mathbf{x}) = \sum_{j=1}^{M} c_j e^{i\mathbf{y}_j \cdot \mathbf{x}} \quad \text{for } \mathbf{x} \in \mathbb{R}^2, \tag{1}$$

i.e., the pairwise distinct frequency vectors $\mathbf{y}_j \in \mathbb{R}^2$ and coefficients $c_j \in \mathbb{C} \setminus \{0\}$.

Several methods for parameter estimation of bivariate exponential sums rely on gridded data [2, 6, 7, 12, 15]. Quite recently, a new algorithm has been proposed, where only a few samples are needed [8, 10]. The key idea in this new approach is to first apply a univariate parameter estimation along specific lines in the plane, before the resulting information are combined to obtain estimations for the frequency vectors $\mathbf{y}_j$ in (1).

This contribution follows along the lines of our ideas in our previous work [3], where the outline of this paper is briefly as follows. In Section 2, we prove that M frequencies $\mathbf{y}_j$ in (1), with arbitrary non-zero coefficients c_j, are uniquely determined already by their projections onto $M+1$ non-parallel lines. Moreover, we discuss the reconstruction of the frequencies from their projections, where special emphasize is placed on the critical issue of cancellation. This then leads us, in Section 3, to parameter estimation of bivariate exponential sums sampled along scattered lines. In Section 4, we show how to recover linear combinations of shifted basic functions from samples in the Fourier domain. For the purpose of illustration, we finally provide one numerical example in Section 5.

In the following of this paper, we use bold letters to denote vectors. Moreover, for $\mathbf{x} \in \mathbb{R}^n$ we let $\|\mathbf{x}\|_0$ denote the number of non-zero components of $\mathbf{x}$. The number of elements in a finite set X is denoted by $|X|$, and we let $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Finally, ∂S denotes the boundary of a set $S \subset \mathbb{R}^d$.

2 Projection of Frequency Vectors onto Scattered Lines

We denote the set of all bivariate exponential sums by

$$\mathcal{E} = \left\{ \sum_{j=1}^{M} c_j e^{i\mathbf{y}_j \cdot \mathbf{x}} : M \in \mathbb{N},\ c_j \in \mathbb{C}^*,\ \mathbf{y}_j \in \mathbb{R}^2 \text{ pairwise distinct} \right\}$$

where we call M the *order* of $f \in \mathcal{E}$ in (1). We first remark that the restriction $f|_\ell$ of any $f \in \mathcal{E}$ to a line $\ell \subset \mathbb{R}^2$ gives a univariate exponential sum. To see this, we represent any line $\ell \subset \mathbb{R}^2$, for some vector $\mathbf{v} \in \mathbb{R}^2$ and scalar $b \in \mathbb{R}$, as

$$\ell \equiv \ell_{\mathbf{v},b} = \{\lambda \mathbf{v} + b\boldsymbol{\eta} : \lambda \in \mathbb{R}\} \subset \mathbb{R}^2,$$

where we assume the orthogonality relation $\mathbf{v} \perp \boldsymbol{\eta}$. In this case,

$$f|_\ell(\lambda) = \sum_{j=1}^{M} c_j e^{i\mathbf{y}_j \cdot (\lambda \mathbf{v} + b\boldsymbol{\eta})} = \sum_{j=1}^{M} c_j e^{ib\mathbf{y}_j \cdot \boldsymbol{\eta}} e^{i\lambda \mathbf{y}_j \cdot \mathbf{v}} = \sum_{j=1}^{M_\ell} c_j^\ell e^{i\lambda y_j^\ell}, \tag{2}$$

where we arrange the representation of $f|_\ell$ on the right hand side of (2), such that the values $y_j^\ell = \mathbf{y}_j \cdot \mathbf{v}$ from the frequency projections onto ℓ are pairwise distinct. Moreover, we assume non-vanishing coefficients $c_j^\ell \in \mathbb{C}^*$, so that the order $M_\ell \leq M$ of $f|_\ell$ is minimal.

Now we consider sampling $f|_\ell$, where we take at least $2M$ equidistant samples. For simplicity, we assume $\|\mathbf{y}_j\|_2 < \pi$ and chose $h = 1$ for the sampling distance. By the application of any suitable parameter estimation method for *univariate* exponential sums, e.g. ESPIRIT [13], we can determine the values

$$(y_j^\ell, c_j^\ell) \qquad \text{for } j = 1, \ldots, M_\ell \tag{3}$$

from the taken samples. Therefore, it remains to recover the unknown parameters

$$(\mathbf{y}_j, c_j) \qquad \text{for } j = 1, \ldots, M. \tag{4}$$

of f in (1) from the parameters (3) of *univariate* exponential sums $f|_\ell$ for several choices of ℓ.

To reformulate this reconstruction problem, we let $X = \{\mathbf{y}_j : j = 1, \ldots, M\}$ and $\mathbf{c} = (c_1, \ldots, c_M)^T \in \mathbb{C}^M$, to associate the data (4) with the *weight function* $w_{X,\mathbf{c}} : \mathbb{R}^2 \to \mathbb{C}$ satisfying

$$w \equiv w_{X,\mathbf{c}}(\mathbf{x}) = \begin{cases} c_j & \text{for } \mathbf{x} = \mathbf{y}_j, \\ 0 & \text{for } \mathbf{x} \in \mathbb{R}^2 \setminus X, \end{cases} \tag{5}$$

so that the support of $w_{X,\mathbf{c}}$ is $\text{supp}(w_{X,\mathbf{c}}) = X$. This allows us to define, for any line $\ell_{\mathbf{v},b} \subset \mathbb{R}^2$, the projection $w_{\mathbf{v},b} : \mathbb{R} \to \mathbb{C}$ of $w_{X,\mathbf{c}}$ onto $\ell_{\mathbf{v},b}$ by

$$w_{\mathbf{v},b}(x) = \sum_{\substack{\mathbf{y} \in X \\ \mathbf{v} \cdot \mathbf{y} = x}} w(\mathbf{y}) e^{ib\mathbf{y} \cdot \boldsymbol{\eta}}, \tag{6}$$

where w denotes the weight function in (5). Therefore, the reconstruction problem for the parameters in (4) boils down to reconstructing $w_{X,\mathbf{c}}$ from a finite number of projections $w_{\mathbf{v}_k,b_k}$, for $k = 1, \ldots, L$.

In the special case, where $b_k = 0$, for all $k = 1, \ldots, L$, and $c_j \in \mathbb{R}_{>0}$, for all $j = 1, \ldots, M$, this leads us to a classical reconstruction problem, as this was studied in [11]. But our situation is somewhat different from that in [11]. To further explain this, note that for $\mathbf{v} \cdot X = \{\mathbf{v} \cdot \mathbf{x} : \mathbf{x} \in X\}$ we have the inequalities

$$|\text{supp}(w_{\mathbf{v},b})| \leq |\mathbf{v} \cdot X| \leq |X|. \tag{7}$$

If the second inequality in (7) is strict, then some points in X must feature the same projection onto $\ell_{\mathbf{v},b}$. In this case, their weights are added. If the sum of their weights is zero, then the first inequality in (7) is also strict. But if equality holds for the second inequality in (7), then so for the first inequality in (7). We remark that for the special case considered in [11] the first inequality in (7) is an equality, i.e., $|\mathrm{supp}(w_{\mathbf{v},b})| = |\mathbf{v} \cdot X|$.

Now note that $\mathbf{v} \mapsto w_{\mathbf{v},b}$ is a linear mapping, i.e., we have

$$(w + t\tilde{w})_{\mathbf{v},b} = w_{\mathbf{v},b} + t\tilde{w}_{\mathbf{v},b} \qquad \text{for any } t \in \mathbb{R} \text{ and } \mathbf{v} \in \mathbb{R}^2.$$

Moreover, for the scaling of vector $\mathbf{v}$, we immediately find the identity

$$w_{t\mathbf{v},b}(x) = \sum_{\substack{\mathbf{y}\in X\\ t\mathbf{v}\cdot\mathbf{y}=x}} w(\mathbf{y})e^{ib\mathbf{y}\cdot\boldsymbol{\eta}} = \sum_{\substack{\mathbf{y}\in X\\ \mathbf{v}\cdot\mathbf{y}=x/t}} w(\mathbf{y})e^{ib\mathbf{y}\cdot\boldsymbol{\eta}} = w_{\mathbf{v},b}(x/t) \qquad \text{for } t \neq 0.$$

In this setting, the result of Renyi in [11] can be stated as follows.

Theorem 2.1 (Renyi, 1952). *For $M+1$ scalar values $b_1, \dots, b_{M+1} \in \mathbb{R}$ and pairwise linearly independent vectors $\mathbf{v}_1, \dots, \mathbf{v}_{M+1} \in \mathbb{R}^2$, let their projections $w_{\mathbf{v}_j,b_j}$ satisfy*

$$\mathrm{supp}(w_{\mathbf{v}_j,b_j}) = \mathbf{v}_j \cdot X \qquad \textit{for all } j = 1, \dots, M+1.$$

Then, $w_{X,\mathbf{c}}$ is uniquely determined by the $M+1$ projections $w_{\mathbf{v}_1,b_1}, \dots, w_{\mathbf{v}_{M+1},b_{M+1}}$.

Proof. We follow an idea of Heppes [5]. To this end, we consider the point set

$$\tilde{X} = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{v}_j \cdot \mathbf{x} \in \mathrm{supp}(w_{\mathbf{v}_j,b_j}) \text{ for all } j = 1, \dots, M+1\} \subset \mathbb{R}^2.$$

Since $\mathrm{supp}(w_{\mathbf{v}_j,b_j}) = \mathbf{v}_j \cdot X$, for $j = 1, \dots, M+1$, we have the inclusion $X \subset \tilde{X}$. To show the inclusion $\tilde{X} \subset X$, let $\mathbf{x} \in \tilde{X}$ be an arbitrary point in $\tilde{X}$. By definition, we find, for each $j = 1, \dots, M+1$, one $\mathbf{x}_j \in X$ satisfying $\mathbf{v}_j \cdot \mathbf{x} = \mathbf{v}_j \cdot \mathbf{x}_j$. As we have picked up $M+1$ points $\mathbf{x}_j$, there is, by the pigeon hole principle, at least one pair of indices $j \neq k$ with $\mathbf{x}_j = \mathbf{x}_k$. But in this case we have

$$\mathbf{v}_j \cdot \mathbf{x} = \mathbf{v}_j \cdot \mathbf{x}_j \quad \text{and} \quad \mathbf{v}_k \cdot \mathbf{x} = \mathbf{v}_k \cdot \mathbf{x}_k = \mathbf{v}_k \cdot \mathbf{x}_j.$$

Since $\mathbf{v}_j, \mathbf{v}_k$ are linearly independent, this implies $\mathbf{x} = \mathbf{x}_j \in X$, and so $\tilde{X} \subset X$. In conclusion, the set $X = \mathrm{supp}(w_{X,\mathbf{c}})$ is uniquely determined by $\tilde{X} = X$.

It remains to recover the weights $\mathbf{c}$ of $w_{X,\mathbf{c}}$ under constraints (5). To this end, note that any $\mathbf{y}_k \in X$ yields $M+1$ projections $\mathbf{v}_j \cdot \mathbf{y}_k \in \mathrm{supp}(w_{\mathbf{v}_j,b_j})$ in $\tilde{X}$. But since there are only $M-1$ points in $X \setminus \{\mathbf{y}_k\}$, there must be at least one vector $\mathbf{v}_j$ satisfying $w_{\mathbf{v}_j,b_j}(\mathbf{v}_j \cdot \mathbf{y}_k) = w_{X,\mathbf{c}}(\mathbf{y}_k) = c_k$. This way, the weights $\mathbf{c}$ of $w_{X,\mathbf{c}}$ are uniquely determined. □

We can conclude that, under the assumptions of Theorem 2.1, any weight function $w_{X,\mathbf{c}} : \mathbb{R}^2 \to \mathbb{C}$ of order M satisfying (5) can uniquely be recovered from $M+1$ projections $w_{\mathbf{v},b}$ of the form (6). As already pointed out in [11], the lower bound $M+1$ for the required number of projections $w_{\mathbf{v},b}$, as given by

Theorem 2.1, is sharp. To see this, let us make one simple example with only M projections. To this end, we regard the $2M$ vertices $\mathbf{x}_j$ of a regular $2M$ polygon $P_{2M} \subset \mathbb{R}^2$, here labelled counterclockwise with indices $j = 1, \ldots, 2M$. If we let $b_j = 0$ and if we denote by $\mathbf{v}_j$ the $2M$ outer normals of P_{2M}, then the M vertices with even indices, $X_e = \{\mathbf{x}_{2j} : j = 1, \ldots, M\}$, have the same projections $w_{\mathbf{v}_j,0}$, $j = 1, \ldots, M$, as the M vertices with odd indices, $X_o = \{\mathbf{x}_{2j-1} : j = 1, \ldots, M\}$.

Further examples are given in [11]. Whether or not the lower bound $M + 1$ for the number of projections, from Theorem 2.1, continues to be sharp for the more general situation, where arbitrary sets of lines are allowed, is still an open problem. Nevertheless, we can show that for any finite set of lines (i.e., finite set of projections), there are at least two different weight functions with equal projections.

Lemma 2.2. *Let $\{\ell_{\mathbf{v}_1,b_1}, \ldots, \ell_{\mathbf{v}_L,b_L}\} \subset \mathbb{R}^2$ be a set of L pairwise distinct lines. Then there is a non-vanishing weight function $w : \mathbb{R}^2 \to \mathbb{C}$, $w \not\equiv 0$, satisfying $w_{\mathbf{v}_j,b_j} = 0$ for all $j = 1, \ldots, L$.*

Proof. We define the convolution of two weight functions $w, \tilde{w} : \mathbb{R}^2 \to \mathbb{C}$ by

$$(w * \tilde{w})(\mathbf{x}) = \sum_{\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}} w(\mathbf{x}_1)\tilde{w}(\mathbf{x}_2). \tag{8}$$

Note that $w * \tilde{w} \equiv 0$ implies $w \equiv 0$ or $\tilde{w} \equiv 0$. Indeed, choosing $\mathbf{x}_1 \in \operatorname{supp}(w)$ and $\mathbf{x}_2 \in \operatorname{supp}(\tilde{w})$ with the largest first component (if multiple choices are available, we choose the vector with the largest second component), respectively, we see that the equation $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}$ has one unique solution with vectors in the support of w and $\tilde{w}$. Hence, $\mathbf{x} \in \operatorname{supp}(w * \tilde{w})$. Further, it is clear that for any pair $w, \tilde{w}$ of weight functions, the sum in (8) is finite. Moreover, the set of weight functions, being equipped with the convolution in (8), is an algebra. The projection of a convolution between two weight functions is given as

$$(w * \tilde{w})_{\mathbf{v},b}(x) = \sum_{(\mathbf{x}_1 + \mathbf{x}_2)\cdot \mathbf{v} = x} e^{ib(\mathbf{x}_1 + \mathbf{x}_2)\cdot \boldsymbol{\eta}} w(\mathbf{x}_1)\tilde{w}(\mathbf{x}_2).$$

We now show that the set of weight functions

$$I_{\mathbf{v},b} = \{w : \mathbb{R}^2 \to \mathbb{C} : w_{\mathbf{v},b} = 0\}$$

is an ideal. In fact, for any $w \in I_{\mathbf{v},b}$ and an arbitrary weight function $\tilde{w}$ we have

$$\begin{aligned}
(w * \tilde{w})_{\mathbf{v},b}(x) &= \sum_{\mathbf{x}_1 \in \operatorname{supp}(\tilde{w})} e^{ib\mathbf{x}_1\cdot\boldsymbol{\eta}} \tilde{w}(\mathbf{x}_1) \sum_{\substack{\mathbf{x}_2 \in \operatorname{supp}(w) \\ \mathbf{x}_2\cdot\mathbf{v} = x - \mathbf{x}_1\cdot\mathbf{v}}} e^{ib\mathbf{x}_2\cdot\boldsymbol{\eta}} w(\mathbf{x}_2) \\
&= \sum_{\mathbf{x}_1 \in \operatorname{supp}(\tilde{w})} e^{ib\mathbf{x}_1\cdot\boldsymbol{\eta}} \tilde{w}(\mathbf{x}_1) w_{\mathbf{v},b}(x - \mathbf{x}_1 \cdot \mathbf{v}) = 0.
\end{aligned}$$

Note that $I_{\mathbf{v},b} \neq \{0\}$, since for instance the weight w satisfying

$$w(0) = 1, \quad w(\boldsymbol{\eta}) = -1, \text{ and } w(\mathbf{x}) = 0 \text{ otherwise}$$

is contained in $I_{\mathbf{v},b}$. But then

$$I_{\mathbf{v}_1,b_1} * \cdots * I_{\mathbf{v}_L,b_L} \subset \bigcap_{j=1}^{L} I_{\mathbf{v}_j,b_j}$$

and so there is one non-trivial weight w satisfying $w_{\mathbf{v}_j,b_j} = 0$ for $j = 1, \ldots, L$. □

We can conclude that it is not possible to uniquely recover any weight function w from an a priori fixed finite number of projections (i.e., from a finite number of lines). This disproves a conjecture in [8].

Note that the proof of Theorem 2.1 is constructive, where we can determine X from $\tilde{X}$. In the more general situation, where $|\mathrm{supp}(w_{\mathbf{v}_j,b_j})| < |\mathbf{v}_j \cdot X|$, the set $\tilde{X}$ does not contain X. In this case, X cannot be determined from $\tilde{X}$. Nevertheless, we can show that $M+1$ projections are sufficient to uniquely determine $w \equiv w_{X,\mathbf{c}}$.

Theorem 2.3. *For a finite set $X = \{\mathbf{y}_j : j = 1, \ldots, M\}$, let $w_{X,\mathbf{c}}$ denote a weight function satisfying (5). Moreover, let $b_1, \ldots, b_{M+1} \in \mathbb{R}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_{M+1} \in \mathbb{R}^2$ be pairwise linearly independent vectors. Then, $w_{X,\mathbf{c}}$ is uniquely determined by the $M+1$ projections $w_{\mathbf{v}_1,b_1}, \ldots, w_{\mathbf{v}_{M+1},b_{M+1}}$ among all weight functions with support size smaller or equal to M.*

Proof. Let $\tilde{w}$ be a different weight function with support size smaller or equal to M satisfying

$$w_{\mathbf{v}_j,b_j} = \tilde{w}_{\mathbf{v}_j,b_j} \qquad \text{for } j = 1, \ldots, M+1,$$

i.e., $(w - \tilde{w})_{\mathbf{v}_j,b_j} \equiv 0$ for $j = 1, \ldots, M+1$. In the following of this proof, we show $|\mathrm{supp}(w - \tilde{w})| \geq 2M + 2$, which is a contradiction to $|\mathrm{supp}(w)|, |\mathrm{supp}(\tilde{w})| \leq M$.

To this end, denote $\mathrm{supp}(w - \tilde{w})$ by Y and note that for any combination of $\mathbf{x} \in Y$ and $\mathbf{v}_j$ there must be one $\mathbf{y} \in Y$, $\mathbf{y} \neq \mathbf{x}$, satisfying $\mathbf{x} \cdot \mathbf{v}_j = \mathbf{y} \cdot \mathbf{v}_j$. Now we consider for $\mathbf{v}_j$ the strip

$$S_j = \{\mathbf{x} \in \mathbb{R}^2 : \min(\mathbf{v}_j \cdot Y) \leq \mathbf{x} \cdot \mathbf{v}_j \leq \max(\mathbf{v}_j \cdot Y)\} \subset \mathbb{R}^2.$$

Note that each of the connected components of the boundary of S_j is a line containing at least two elements from Y. Further note that Y is contained in S_j, i.e., $Y \subset S_j$, for all $j = 1, \ldots, M+1$, and so the intersection

$$P = \bigcap_{j=1}^{M+1} S_j$$

is a convex polygon containing Y, i.e., $Y \subset P$.

Now the polygon P has $2M+2$ edges. Indeed, for any index $k \in \{1,\dots,M+1\}$, the boundary lines of S_k have a non-empty intersection with the polygon

$$P_k = \bigcap_{\substack{j=1 \\ j\neq k}}^{M+1} S_j.$$

In fact, this intersection $\partial S_k \cap P_k$ is given by two line segments, containing at least two points from Y that are lying on the boundary of S_k. But these line segments cannot be edges in P_k, since the vectors $\mathbf{v}_j$ were assumed to be pairwise linearly independent. In conclusion, P has $2M+2$ edges, two from each intersection $S_k \cap P_k$, for $k = 1,\dots,M+1$. Moreover, each edge of P contains at least two elements from Y. Therefore, Y has at least $2M+2$ elements.

But this implies $|\mathrm{supp}(w)| = |X| > M$, which completes our proof. □

We remark that, in contrast to Theorem 2.1, the proof of Theorem 2.3 is not constructive, and so the support X of $w_{X,\mathbf{c}}$ cannot be recovered directly. But it is possible to construct a finite candidate set, containing X.

Lemma 2.4. *Let $M \in \mathbb{N}$. Moreover, for $b_1,\dots,b_{M+1} \in \mathbb{R}$ and pairwise linearly independent vectors $\mathbf{v}_1,\dots,\mathbf{v}_{M+1} \in \mathbb{R}^2$, let $w_{\mathbf{v}_j,b_j}$, $j = 1,\dots,M+1$, denote their corresponding projections. Then we have the inclusion*

$$X \subset \tilde{X} = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{v}_j \cdot \mathbf{x} \in \mathrm{supp}(w_{\mathbf{v}_j,b_j}) \text{ for at least two distinct } j\}. \tag{9}$$

Proof. Let $\mathbf{x} \in X$. If $\mathbf{x} \notin \mathrm{supp}(w_{\mathbf{v}_j,b_j})$, for one index $j \in \{1,\dots,M+1\}$, then there must be one $\mathbf{x}_j \in X$ satisfying $\mathbf{x}\cdot\mathbf{v}_j = \mathbf{x}_j \cdot \mathbf{v}_j$. Since we can pick only at most $M-1 = |X \setminus \{\mathbf{x}\}|$ pairwise distinct points $\mathbf{x}_j$ from $X \setminus \{\mathbf{x}\}$, there must be at least two distinct indices $j,k \in \{1,\dots,M+1\}$ satisfying $\mathbf{v}_j \cdot \mathbf{x} \in \mathrm{supp}(w_{\mathbf{v}_j,b_j})$ and $\mathbf{v}_k \cdot \mathbf{x} \in \mathrm{supp}(w_{\mathbf{v}_k,b_k})$. □

If we wish to use Lemma 2.4 to recover X, then we would have to check all subsets of $\tilde{X}$ in (9) of size at most M. Since $|\tilde{X}| \leq \binom{M}{2} M^2$ this is computationally not feasible. Yet it is possible to reduce the size of the candidate set $\tilde{X}$.

Lemma 2.5. *Under the same assumptions and notations as in Lemma 2.4, let the index set $J = \{r_1,...,r_s\} \subset \{1,...,M+1\}$ be chosen such that*

$$|\mathrm{supp}(w_{\mathbf{v}_r,b_r})| \geq M-1 \qquad \text{for all } r \in J.$$

Then we have the inclusion

$$X \subset X_J = \{\mathbf{x} \in \tilde{X} : \mathbf{x}\cdot\mathbf{v}_r \in \mathrm{supp}(w_{\mathbf{v}_r,b_r}) \text{ for all } r \in J\}.$$

Moreover, the set

$$\tilde{X}_J = \{\mathbf{x} \in X_J : \forall j \in \{1, \ldots, M+1\} \setminus J : \\ \mathbf{x} \cdot \mathbf{v}_j \in \operatorname{supp}(w_{\mathbf{v}_j, b_j}) \text{ or } \exists \mathbf{y} \in X_J, \mathbf{x} \neq \mathbf{y}, \text{ with } \mathbf{x} \cdot \mathbf{v}_j = \mathbf{y} \cdot \mathbf{v}_j\}$$

also contains X, i.e., $X \subset \tilde{X}_J$.

Proof. If $|\operatorname{supp}(w_{\mathbf{v}_r, b_r})| \geq M - 1$, then $\mathbf{v}_r \cdot X = \operatorname{supp}(w_{\mathbf{v}_r, b_r})$, and so $X \subset X_J$. If $\mathbf{x} \in X_J$ is not contained in $\operatorname{supp}(w_{\mathbf{v}_j, b_j})$ but in X, then there is one $\mathbf{y} \in X \setminus \{\mathbf{x}\}$ with the same projection onto $\ell_{\mathbf{v}_j, b_j}$. This yields $X \subset \tilde{X}_J$. □

Unfortunately, there are cases where $J = \emptyset$ or $|J| = 1$ and $X \subsetneq \tilde{X}_J$. But the reduction of the candidate set, as suggested in Lemma 2.5, works only for $|X| = M$, whereas the other results in this section also hold for $|X| \leq M$.

We conclude this section by referring to a few other possibilities to perform an efficient recovery of X. In [8], an adaptive choice of lines $\ell_{\mathbf{v}_j, b_j}$ is considered. In fact, as shown in [8], if two lines $\ell_{\mathbf{v}_j, b_j}$ with $\operatorname{supp}(w_{\mathbf{v}_j, b_j}) = \mathbf{v}_j \cdot X$, $j = 1, 2$, are known, the projection onto an adaptively chosen third line $\ell_{\mathbf{v}_3, b_3}$ is sufficient to uniquely recover the weight function w, where $\ell_{\mathbf{v}_3, b_3}$ depends on $w_{\mathbf{v}_j, b_j}$, $j = 1, 2$.

The problem to recover a set of vectors from a finite number of projections is also discussed in [1]. We can state one of their results in our setting as follows. For given projections onto $s = \binom{M}{2} + 2$ lines $\ell_{\mathbf{v}_j,\, b_j}$ — with points $\mathbf{v}_j$, $j = 1, \ldots, s$, in general position — there are at least two distinct lines r_1, r_2 with $|\mathbf{v}_{r_j} \cdot X| = M$. Note that this implies $|\operatorname{supp}(w_{\mathbf{v}_{r_j}, b})| = M$.

3 Sampling along Scattered Lines

We now return to the problem of parameter estimation for bivariate exponential sums. The following result is a direct consequence from Theorem 2.3.

Corollary 3.1. *Let $f \in \mathcal{E}$ be a bivariate exponential sum of order M and let $\ell_{\mathbf{v}_1, b_1}, \ldots, \ell_{\mathbf{v}_{M+1}, b_{M+1}}$ be pairwise non-parallel lines. Moreover, let G be a set containing at least $2M$ equidistant sample points along each of the lines $\ell_{\mathbf{v}_j, b_j}$, $j = 1, \ldots, M+1$, taken with step size one. Further, assume that every frequency vector $\mathbf{y}$ of f satisfies $\|\mathbf{y}\|_2 < \pi$. Then, f is uniquely determined among all exponential sums of order at most M by its samples at the points in G.*

To describe a procedure how to recover $f \in \mathcal{E}$ from its restrictions $f|_{\ell_{\mathbf{v}_j, b_j}}$, we reformulate Lemma 2.4 as follows.

Corollary 3.2. *Under the assumptions and notations of Corollary 3.1, let X_j be the frequency set of the restriction $f|_{\ell_{\mathbf{v}_j, b_j}}$, $j = 1, \ldots, M+1$. Moreover, let*

$$\tilde{X} = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} \cdot \mathbf{v}_j \in X_j \text{ for at least two distinct } j\}.$$

Then the optimization problem

$$\begin{aligned} &\min_{\mathbf{c}\in\mathbb{R}^{|\tilde{X}|}} \|\mathbf{c}\|_0 \\ &\text{subject to} \sum_{\mathbf{x}\in\tilde{X}} c_{\mathbf{x}} e^{i\mathbf{x}\cdot\mathbf{w}} = f(\mathbf{w}) \text{ for all } \mathbf{w}\in G \end{aligned} \tag{10}$$

has a unique solution corresponding to f.

Proof. First note that the optimization problem (10) has a solution satisfying $\|\mathbf{c}\|_0 = M$, corresponding to f. Any other solution with $\|\tilde{c}\|_0 \le M$ gives rise to an exponential sum $\tilde{f}$ of order at most M with $f(\mathbf{w}) = \tilde{f}(\mathbf{w})$ for all $\mathbf{w} \in G$. But this implies $f = \tilde{f}$, on Corollary 3.1. □

Unfortunately, the optimization problem (10) is NP-hard, see [4]. Moreover, we cannot expect its system matrix to satisfy the restricted isometric property, which would allow us to consider the convex relaxation with replacing $\|\mathbf{c}\|_0$ by $\|\mathbf{c}\|_1$. Also, $\tilde{X}$ is very large. If we knew M exactly, we could possible reduce $|\tilde{X}|$ quite significantly by the application of Lemma 2.5.

We close this section by providing a numerical algorithm to recover any bivariate exponential sum $f \in \mathcal{E}$ of order $M \in \mathbb{N}$. The following algorithm can be viewed as a modification of the *Sparse Approximate Prony Method* in [10].

Algorithm 3.3. *Let $f \in \mathcal{E}$ be an exponential sum of order $M \in \mathbb{N}$. Choose $\varepsilon, \tilde{\varepsilon} > 0$ and $\ell_{\mathbf{v}_1,b_1}, \ldots, \ell_{\mathbf{v}_L,b_L}$ for $L \ge M$. Take $2L$ samples from f along each line $\ell_{\mathbf{v}_j,b_j}$, at equidistant sample points and for step size one, respectively.*

1. *Apply algorithm ESPRIT [13] along each line $\ell_{\mathbf{v}_j,b_j}$, with rank estimation parameter ε. Let X_j be the frequencies observed on the j-th line $\ell_{\mathbf{v}_j,b_j}$.*
2. *Let $J = \{r_1, ..., r_s\}$ be the index set with $|X_j| \ge L - 1$. Build $\tilde{X}_J$ as in Lemma 2.5. Eliminate point pairs with Euclidean distance smaller than $\tilde{\varepsilon}$.*
3. *Solve*

$$\begin{aligned} &\min_{\mathbf{c}\in\mathbb{R}^{|\tilde{X}_J|}} \|\mathbf{c}\|_0 \\ &\text{subject to} \sum_{\mathbf{x}\in\tilde{X}_J} c_{\mathbf{x}} e^{i\mathbf{x}\cdot\mathbf{w}} = f(\mathbf{w}) \text{ for all } \mathbf{w}\in G \end{aligned} \tag{11}$$

We remark that the tolerance ε in Algorithm 3.3 should to be chosen with respect to the noise level, whereas $\tilde{\varepsilon}$ could be adapted to the separation distance of the candidate set $\tilde{X}_J$. If M is known, then we usually have $\tilde{X}_J = X$, in which case the solution of the optimization problem (11) is feasible.

4 Reconstruction from Fourier Data

In this section, we apply the parameter estimation method of the previous section to reconstruct linear combinations of a shifted basic function from samples in the Fourier domain (cf. [8, 16]). To this end, we follow along the lines of our work [3].

We define the (continuous) Fourier transform $\hat{f}$ of a function $f \in L^1(\mathbb{R}^2)$ by

$$\hat{f}(\mathbf{w}) = \frac{1}{2\pi} \int_{\mathbb{R}^d} f(\mathbf{x}) e^{i\mathbf{x}\cdot\mathbf{w}} d\mathbf{x} \qquad \text{for } \mathbf{w} \in \mathbb{R}^2.$$

Let $\Phi : \mathbb{R}^2 \to \mathbb{R}$ be an even function. For f we assume the form

$$f(\mathbf{x}) = \sum_{j=1}^{M} c_j \Phi(\mathbf{x} - \mathbf{x}_j), \tag{12}$$

where $\Phi \in L^1(\mathbb{R}^2)$. The Fourier transform of f is then given by

$$\hat{f}(\mathbf{w}) = \hat{\Phi}(\mathbf{w}) \sum_{j=1}^{M} c_j e^{i\mathbf{w}\cdot\mathbf{x}_j}.$$

Sampling $\hat{f}$ at a finite point set G leads us to a reconstruction problem for bivariate exponential sums, provided that $\hat{\Phi} \not\equiv 0$ on G. As Φ is assumed to be even, $\hat{\Phi}$ is real-valued. If $\hat{\Phi}$ is strictly positive (or strictly negative) on $\mathbb{R}^2$, then we do not need to make any assumptions on G. To further explain this, note that

$$\hat{\Phi}(\mathbf{w}) > 0 \qquad \text{for all } \mathbf{w} \in \mathbb{R}^2$$

implies that Φ is a positive definite function, by Bochner's theorem. Positive definite functions are often considered in approximation theory. Prototypical examples of such functions are the *Gaussians* $\Phi(\mathbf{x}) = e^{-\alpha\|\mathbf{x}\|_2^2}$, for $\alpha > 0$. Their Fourier transform is given by

$$\hat{\Phi}(\mathbf{w}) = \frac{1}{2\alpha} e^{-\|\mathbf{w}\|_2^2/(4\alpha)} > 0.$$

Other examples are the *inverse multiquadrics*

$$\Phi(\mathbf{x}) = \left(1 + \|\mathbf{x}\|_2^2\right)^{\beta} \qquad \text{for } -2 < \beta < 0.$$

Now we can transfer our preceding considerations, in particular Algorithm 3.3, to obtain a reconstruction method for functions f of the form (12).

(1) Sample $\hat{f}$ on enough lines, as described in Algorithm 3.3.

(2) Calculate

$$g(\mathbf{w}) = \frac{\hat{f}(\mathbf{w})}{\hat{\Phi}(\mathbf{w})} = \sum_{j=1}^{M} c_j e^{i\mathbf{w}\cdot\mathbf{x}_j}$$

for all sample points.

(3) Use Algorithm 3.3 to reconstruct the frequencies and coefficients of g. The frequencies are the shift vectors $\mathbf{x}_j \in \mathbb{R}^2$.

Although the above algorithm works fine at the absence of noise, it is not very stable: As we divide by $\hat{\Phi}(\mathbf{w})$, we require $\hat{\Phi}$ to be uniformly bounded away from zero, i.e.,

$$\hat{\Phi}(\mathbf{w}) > C > 0 \tag{13}$$

for some sufficiently large constant C. Otherwise any noise added to $\hat{f}$ will get amplified significantly. But due to the Riemann-Lebesgue lemma, we have

$$\hat{\Phi}(\mathbf{w}) \longrightarrow 0 \quad \text{for } \mathbf{w} \to \infty,$$

for $\Phi \in L^1(\mathbb{R}^2)$. Therefore, (13) can only hold on a bounded set.

5 Numerical Example

For the purpose of illustration, we finally provide one numerical example. Let

$$\begin{pmatrix}\mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \mathbf{y}_4 & \mathbf{y}_5\end{pmatrix} = \begin{pmatrix} 1 & 1.5 & 2 & 2 & 2.8 \\ 1 & 2.7 & 1 & 2.2 & 1.6 \end{pmatrix}$$
$$\mathbf{c} = \begin{pmatrix} 1 & 2 & -1 & 3 & 0.5 \end{pmatrix}$$

We sample along four lines with $b_j = 0$ and $\mathbf{v}_j = (\cos(\varphi_j), \sin(\varphi_j))$, where $\varphi_j = 0, \pi/2, \pi/4, -\pi/4$. This example is numerically critical, as cancellation occurs: The vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ have the same projection onto $\ell_{\mathbf{v}_2,0}$ and their coefficients sum up to zero. Ignoring $\ell_{\mathbf{v}_2,0}$ does not help in this case, as then two additional frequency vectors appear and $X \subsetneq \tilde{X}_J$.

We apply Algorithm 3.3 to this problem with letting $\varepsilon = \tilde{\varepsilon} = 10^{-3}$ and $L = 5$. Thus, the reduction step of Lemma 2.5 is applicable. We decided to take 21 samples along each line. As all coefficients are real-valued and $b_j = 0$ for $j = 1, ..., 5$, we may rely on the identity $f(-\mathbf{x}) = \overline{f(\mathbf{x})}$.

Thus, we can use 41 samples for each univariate method. Further, we have added noise, uniformly and independently distributed in $[-\delta, \delta]$ to each sample. We performed each calculation 100 times and recorded the errors

$$e_{\text{freq}} = \max_{j=1,\dots,5} \|\mathbf{y}_j - \tilde{\mathbf{y}}_j\|_\infty \qquad \text{and} \qquad e_{\text{coef}} = \max_{j=1,\dots,5} \|c_j - \tilde{c}_j\|_\infty$$

Table 1: Results of the Example

δ	e_{freq}	e_{coef}	fails
0	4.9e-15	2.7e-14	0
10^{-6}	3.5e-7	5.1e-7	0
10^{-3}	2.9e-4	4.9e-4	0
10^{-2}	2.6e-3	4.8e-3	10

We averaged the error over the 100 experiments. Any time when the number of frequencies was not estimated correctly, we counted one fail. We observed that fails occur whenever ESPRIT did not estimate the number of frequencies along one of the lines correctly. Our numerical results are shown in Table 1.

References

[1] M.D. Buhmann and A. Pinkus: On a recovery problem. *Annals of Numerical Mathematics* **4**, 1997, 129–142.

[2] E.J. Candès and C. Fernandez-Granda: Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics* **67**(6), 2014, 906–956.

[3] B. Diederichs and A. Iske: Parameter estimation for bivariate exponential sums. *IEEE International Conference Sampling Theory and Applications* (SampTA2015), 2015, 493–497.

[4] S. Foucart and H. Rauhut: *A Mathematical Introduction to Compressive Sensing*. Springer, 2013.

[5] A. Heppes: On the determination of probability distributions of more dimensions by their projections. *Acta Mathematica Hungarica* **7**(3-4), 1956, 403–410.

[6] Y. Hua: Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Transactions on Signal Processing* **40**(9), 1992, 2267–2280.

[7] S. Kunis, T. Peter, T. Roemer, and U. von der Ohe: A multivariate generalization of Prony's method. *Linear Algebra and its Applications* **490**, 2016, 31–47.

[8] G. Plonka and M. Wischerhoff: How many Fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing* **42**(1-2), 2013, 117–137.

[9] D. Potts and M. Tasche: Parameter estimation for exponential sums by approximate Prony method. *Signal Processing* **90**(5), 2010, 1631–1642.

[10] D. Potts and M. Tasche: Parameter estimation for multivariate exponential sums. *Electronic Transactions on Numerical Analysis* **40**, 2013, 204–224.

[11] A. Renyi: On projections of probability distributions. *Acta Mathematica Hungarica* **3**(3), 1952, 131–142.

[12] S. Rouquette and M. Najim: Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods. *IEEE Transactions on Signal Processing* **49**(1), 2001, 237–245.

[13] R. Roy and T. Kailath: ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing* **37**(7), 1989, 984–995.

[14] T.K. Sarkar and O. Pereira: Using the matrix pencil method to estimate the parameters of a sum of complex exponentials. *IEEE Antennas and Propagation Magazine* **37**(1), 1995, 48–55.

[15] T. Sauer: Prony's method in several variables. *arXiv preprint arXiv:1602.02352*, 2016.

[16] M. Wischerhoff: *Reconstruction of Structured Functions From Sparse Fourier Data*. PhD thesis, Universität Göttingen, 2015.